Research Paper | Computer Science | E-ISSN No : 2454-9916 | Volume : 5 | Issue : 9 | Sep 2019

DETECTING DOS ATTACKS IN THE INTERNET OF THINGS (IOT) USING INTELLIGENT NEURAL NETWORK

Farzane Mohamady[1,*] | Pouya Asadi[2] | Akbar Morshed Aski[3]

[1]Master of Software engineering, Department of Computer, Islamic Azad University Varamin Branch, Varamin, Iran.
[2]Ph.D. of Software engineering, Department of Computer, Islamic Azad University Varamin Branch, Varamin, Iran.
[3]Master of Software engineering, Department of Computer, Islamic Azad University Varamin Branch, Varamin, Iran.

ABSTRACT

The ICT is increasingly used in the last decade. One of the greatest challenges in this area is system and information security. This security is threatened by cyberattacks on networks. New techniques have been devised for the field by expanding research into network security. Among the most popular techniques, various data mining algorithms can be noted that are used either alone or in combination for intrusion detection. In this study, neural network technique was used for intrusion detection. The results of simulation of this technique on KDD dataset and MATLAB software indicated high precision of 98%. Comparison of this technique with other techniques such as CNN and BiLSTM suggested that these techniques had a less precision than the proposed one. In addition, the technique had a low error rate.

KEYWORDS: intrusion detection, cyberattacks, neural networks, data mining.

INTRODUCTION:

Internet of things or IoT refers to a set of devices used in a field and connected together via the Internet. One of the most challenging issues in IoT, which has a large number of devices connected to the Internet, is the security and privacy of these networks. Thus, the inherent vulnerabilities in cyberspace and the growing trend of migration from the traditional world to this space increase the risk of IT-based systems, which are vital to the economies of nations. Authentication and access control techniques are as central elements to address the issues of security and privacy. These techniques prevent unauthorized users from accessing the network and prohibit unauthorized access by authorized users. For this purpose, intrusion detection techniques are used in IoT (Chen et al., 2010). The purpose of intrusion detection systems (IDSs) to detect and identify different kinds of intruders to the network (Elhag, 2015). Intrusion detection systems (IDSs) are generally divided into two groups: anomaly detection systems and misuse detection systems (MDSs). In MDSs, instructions are identified based on parameters system weaknesses and known signatures. However, it is not possible to detect novel or unconscious attacks by these systems (Elhag, 2015). A very important point about intrusion detection systems is that in some cases attacks are mis-detected and so the number of notifications to the user increases (Nandeshwar and Bijwe., 2015).

Using intrusion detection systems in the area of computer and network security, users can be alerted of network risks. Once these alerts are received, the system will take appropriate measures by other entities. Two important tools in this area are Firewall and IDS. There are differences between the two tools as follows (Meena and Choudhary, 2017):

1. In the firewall, it is not possible to detect intrusion from within the network, but in IDS systems all network traffic can be reviewed.

2. The firewall is mainly used to prevent unauthorized users from accessing the internal network, but IDS systems are used to detect attacks on networks.

These attacks are highly dangerous and threaten the security of the system. As a result, systems decided to use security systems, firewalls, and intrusion detection systems to counter these attacks (Nandeshwar and Bijwe, 2015).

Different techniques have been proposed to detect attacks on the IoT and on computer networks, each of which has its advantages and disadvantages. None of these techniques are 100% precise. All areas of using intelligent neural networks to solve problems were well precise (Kumar and Gohil, 2015). In intrusion detection systems based on analytical technologies, detection is performed in two different ways (Buczak, and Guven, 2015): 1. Anomaly detection 2. Signature-based detection. In the anomaly detection-based model, a normal profile with data on system activities is made and used to identify the pattern of activities. In the signature-based technique, the IDSs examine the network and system activity using the pattern matching algorithm. This technique considers patterns of known attacks to detect attacks (Nskh et al., 2016). Signature-based systems have a very low false positive (FP) rate. This means in these systems, the rate of non-intrusive errors that are mis-detected is very low. Additionally, these systems are unable to detect novel and vague attacks. On the other hand, anomaly-based systems produce a lot of FP. It is therefore suggested that both techniques are used in an IDS (Ranjan and Sahoo, 2014).

In their study, Aljawarneh et al (2018) compared the precision of some of the techniques used for intrusion detection. The results indicated that the j48 method was the most precise and the Navie Bayesian method was the least precise (Kumar and Gohil, 2015). In 2019, to improve precision, the normal data od system were first divided into categories based on the similarity detection performed by the user and then, the anomalies or intrusions into the system were detected using the HAD Hierarchical Algorithm. The precision of this technique was 99.76% (Kim et al., 2014). In one study (2019), a mechanism for punishing nodes was used to reduce the rate of false notification. The results suggested that this technique greatly reduced the rate of false notification in attack detection. The precision of this technique was 97.62% (Yahalom et al., 2019). In 2019, the feature detection technique was found to be effective for intrusion detection. For this purpose, features were first detected and clustered using feature grouping based on linear correlation coefficient (FGLCC) filtering method. Finally, decisions were made on normal data or intrusion using the decision tree methodology. The final precision of this technique was 95.03% (Guo et al., 2019). In 2019, features were initially reduced using feature detection. The results indicated that in this technique, the support vector machine (SVM) algorithm was the most precise (Mohammadi et al., 2019). In this study (1), the use of lazy learning approaches was suggested to improve the overall performance of IDS. The results of this study openly introduced periodic algorithms as a perfect solution for intrusion detection in real networks (Dutt et al., 2018).

According to previous studies, in this study, a method for detecting DOS attacks in the system was introduced. In this method, intelligent neural network method was used to classify the data into two categories: normal and abnormal. To this end, feedback-based neural network was used after the data was cleaned and standardized. The proposed method is expected to improve the precision of attack detection in system.

METHODOLOGY:

In this method, the designed neural network has two types of learning. The first type is supervised learning using the training data.

$$D=\{(x_i,y_i)\}_{i=1}^{N} \quad (1)$$

In this equation, d is the training set and N is the number of training examples. It is also assumed that yi is a classification of variables from a similar set such as yiε {1, ... , c}. The second type is unsupervised learning, in which the unlabeled data d = {xi} i = 1^N are entered and then the neural network finds patterns among them. To do this, the MLP architecture with three layers of neural network, as shown in Figure 1, was used. This technique has 6 inputs, three middle layers and one output. This network is trained using forward and backward learning algorithms.

The first phase: Forward learning algorithm:

S_i is the number of units containing bias. So, we will have a network with the following parameters:

$$(w,b) = (w^{(1)}, b^{(1)}, w^{(2)}, b^{(2)}) \quad (2)$$

In Equation 2, the parameter w represents the parameters associated with the connections of the unit j of the layer l and the unit i of the layer l + 1. Also, b is the bias associated with the unit i of the layer l + 1. This model is shown mathematically in equations 1 to 5 (Kumar and Gohil, 2015).

$$a_1^{(2)} = f(w_{11}^{(1)} x_1 + w_{12}^{(1)} x_2 + \ldots + w_{16}^{(1)} x_6 + b_1^{(1)}) \quad (3)$$

$$a_2^{(2)} = f(w_{21}^{(1)} x_1 + w_{22}^{(1)} x_2 + \ldots + w_{26}^{(1)} x_6 + b_2^{(1)}) \quad (4)$$

$$a_3^{(2)} = f(w_{31}^{(1)} x_1 + w_{32}^{(1)} x_2 + \ldots + w_{36}^{(1)} x_6 + b_3^{(1)}) \quad (5)$$

$$h_{w,b}(x) = f(w_{11}^{(2)} a_1^{(2)} + w_{12}^{(2)} a_2^{(2)} + w_{13}^{(2)} a_3^{(2)} + b_1^{(2)}) \quad (6)$$

The second phase: Feedback or backward layer:

Feedback mode in this network is according to Equation 7:

$$\delta_i^{(l_2)} = \frac{\partial}{\partial z_i^{(l_2)}} \frac{1}{2} \left\| y - h_{w,b}(x) \right\|^2 = -(y_i - a_i^{(l_2)}) \cdot f'(z_i^{(l_2)}) \quad (7)$$

Finally, the error rate can be measured for evaluation.

$$\delta_i^{(l)} = \left(\sum_j w_{ji}^{(l)} \delta_j^{(l+1)}\right) f'(z_i^{(l_2)}) \quad (8)$$

The diagram of the proposed method can be seen in Figure 2.

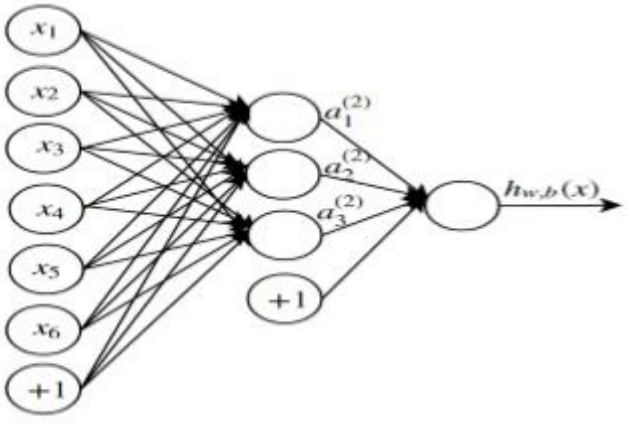

Figure 1: Architecture of the proposed method

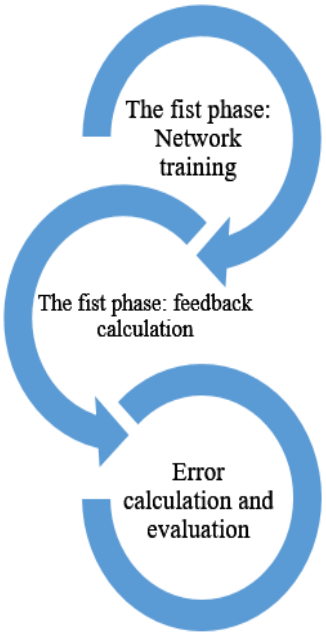

Figure 2: Diagram of the proposed method

RESULTS AND DISCUSSION:

In this method, the dataset was used for data entry. In the first step, the KDD dataset was used. The number of records in this dataset was over 5 million. For this reason, a database containing 10% of the records of this dataset was used. The dataset was divided into two parts, d1 training data and d2 experimental data. KDD CUP 99 dataset was used to produce training and experimental datasets. This dataset contained more than 4 million records was is too large for simulation. So according to (Salo, 2019), the training dataset consists of 60593 normal records, 49115 records of DOS attack data, 1917 records of probe attack, 899 R2Ls and 26 data records for the U2R that were randomly selected from the original dataset. This dataset provided 41 features for each input data. Of information classes, 5 normal classes including Dos, Probe, R2L and U2R attack were selected as output data. Intrusion detection system evaluation criteria include recall, precision and FM, which were calculated as follows:

$$\text{Recall} = \frac{\text{Tp}}{\text{TP} + \text{F}} \quad (9)$$

$$\text{Precision} = \frac{\text{Tp}}{\text{TP} + \text{FP}} \quad (10)$$

$$\text{FM} = \frac{2.\text{Recall}.\text{Precision}}{\text{Recall} + \text{Precision}} \quad (11)$$

Initial values and values after the implementation of the method can be seen in Tables 1 and 2, and confusion matrix values for both data series are shown in Tables 3 and 4.

Table 1: Basic settings of the MLP implementation

Parameter	Value
Function name	MLP
Mean squared error (MSE)	983.6248
Number of features	38
Number of experimental samples	6492
Number of training samples	22683
Number of iterations	1500
Total number of samples	29175

Table 2: Results of MLP implementation

Size	Training data	Experimental data
accuracy	98.9243	72.9051
Specificity	0.98701	0.78821
Recall	0.98701	0.78821
F1	0.98882	0.72683
gmean	0.98694	0.76842

Table 3. Confusion matrix for training data

	Predict no	Predict yes
Actual no	13436	13
Actual yes	231	9003

Table 3. Confusion matrix for experimental data

	Predict no	Predict yes
Actual no	2074	78
Actual yes	1681	2959

Comparison with other techniques:

To compare this technique with other ones, two techniques of CNN and BiLSTM were selected among neural network techniques. A CNN usually has three layers: the cannulation layer, the collector layer, and the fully connected (FC) layer. The convolution layer is the core block of CNN.

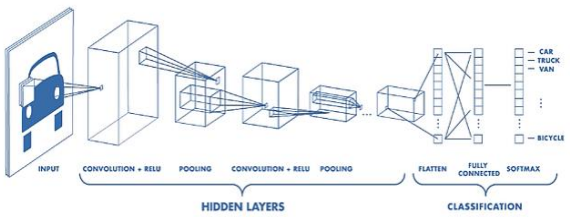

Figure 3: CNN architecture [26]

Long-short term memory networks (commonly called "LSTMs") are a specific type of RNN that can learn long-term dependencies. All neural networks are a chain of iteration modules. In standard RNNs, this iteration module has a very simple structure such as a single layer. Figure 4 shows the LSTM diagram. In this study, LSTM is proposed to predict probability of DDoS attack based on input data.

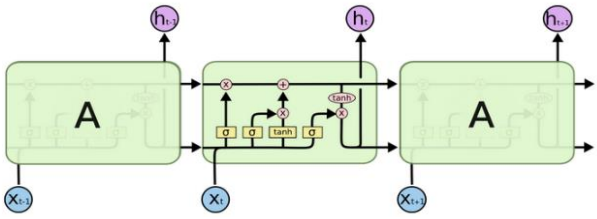

Figure 4: LSTM diagram

The following equations are used in this network:

$$i_t = \delta\left(W_{ix}x_t + W_{im}m_{t-1} + W_{ic}c_{t-1} + b_i\right)$$

$$f_t = \delta\left(W_{f_x}x_t + W_{fm}m_{t-1} + W_{fc}c_{t-1} + b_f\right)$$

$$c_t = f_t \otimes c_{t-1} + i_t \otimes g\left(W_{cx}x_t + W_{cm}m_{t-1} + b_c\right)$$

$$o_t = \delta\left(W_{ox}x_t + W_{om}m_{t-1} + W_{oc}c_t + b_0\right)$$

$$m_t = o_t \otimes h\left(c_t\right)$$

$$y_t = W_{ym}m_t + b_y$$

There is a special type of LSTM called BiLSTM. In this structure, learning takes place one the data once from the beginning to the end and once again from the end to the beginning. In Figures 5 and 6, the structure of LSTM and BiLSTM can be seen.

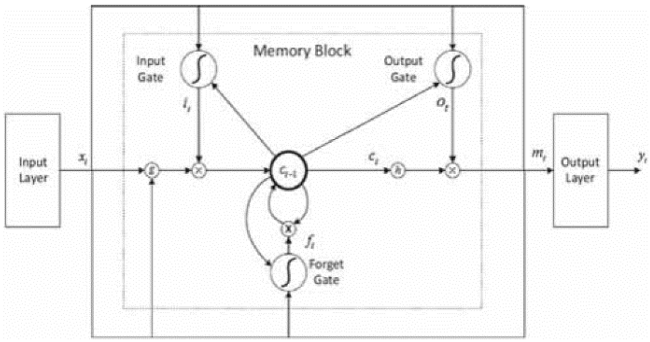

Figure 5: LSTM structure

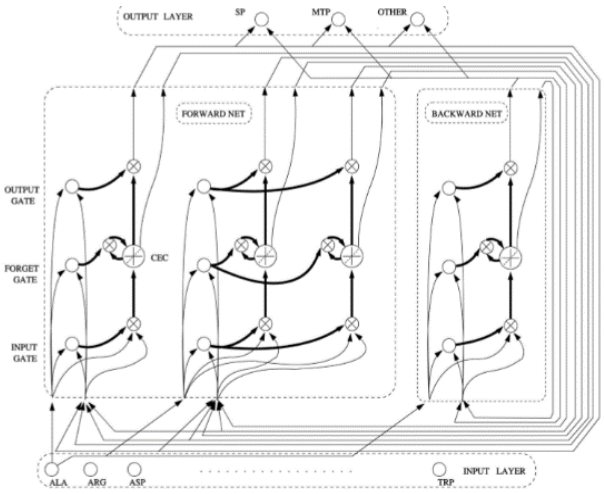

Figure 6: BiLSTM architecture

Of the two techniques mentioned, CNN was selected for implementation. Several criteria can be used to compare these techniques. Precision is one of the most important criteria. A comparison of the precision of the four techniques referred to can be seen in Figure 7 and it is obvious that the MLP technique had more precision. Also, the confusion matrix can be used to compare mentioned techniques. By comparing the error values we can see in which technique the error rate was less. This rate can be seen in Figure 8. As can be seen, the error rate in the MLP was lower than other techniques.

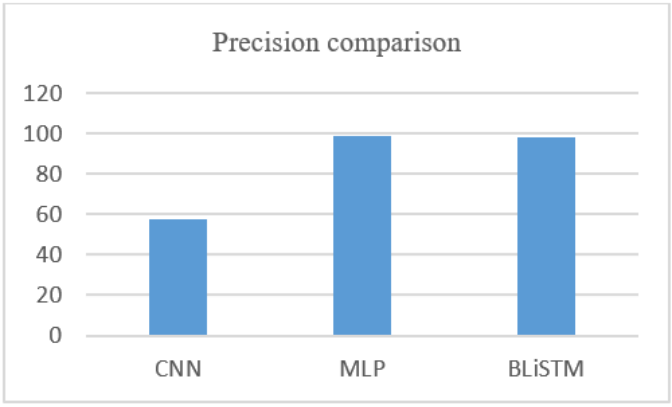

Figure 7: Comparison of precision in the four techniques mentioned

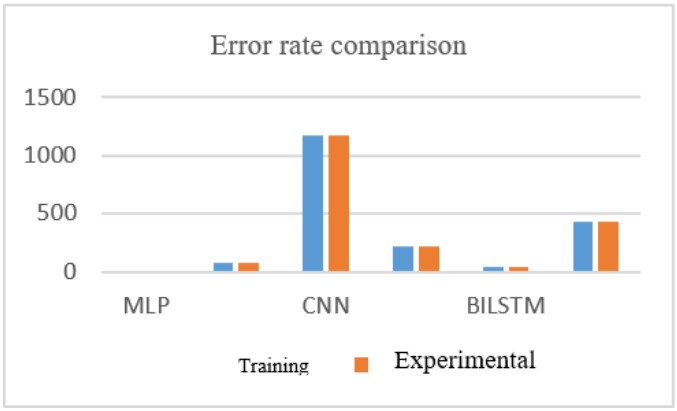

Figure 8. Error rate in categorized data not detected as an attack

CONCLUSION:

Due to increasing number of cyberattacks and the varied use of them, various measures and strategies have been proposed by government and private security agencies. Among the major problems that cause system attacks and threaten system security, the hiring of untrained and unskilled experts, lack of awareness of the impact of a security weakness on the organizational performance, lack of proper budget allocation to address information security, and complete reliance on business tools and products can be mentioned. Today, building a strong security structure can be effective in raising the security factor as well as reducing the risk of possible attacks. One of the steps that provide a secure environment for authorized users to access the services needed in the computing environment is intrusion testing. System testing should be performed at regular intervals to ensure the provider system. Since in each technique we try to improve security in a particular way and each algorithm is more useful in one of these techniques, the use of different algorithms in this field is very different.

As mentioned above, this study introduced a technique for intrusion detection in which the neural network method was used. The results indicated that this technique, following network training and high number of attacks, could show a better performance and reduce the rate of delay in attack detection. This parameter is widely used in many networks. Among the other benefits of this technique are the low number of mis-detected attacks. Given the increasing use of networks in the present era and the need for security in them, a technique must be used that can be establish security. So the proposed technique can be used by all users working in the network and all network administrators. Given the improved results in network attack detection using meta-heuristic algorithms, future researchers are recommended to use a combination of these techniques such as feature selection algorithm and MLP, and feature selection algorithm and RBF network attack detection.

REFERENCES:

1. Buczak, A. L., & Guven, E. (2015). A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications Surveys & Tutorials, 18(2), 1153-1176.
2. Chen, C. M., Chen, Y. L., & Lin, H. C. (2010). An efficient network intrusion detection. Computer Communications, 33(4), 477-484.
3. Dutt, I., Borah, S., Maitra, I. K., Bhowmik, K., Maity, A., & Das, S. (2018). Real-time hybrid intrusion detection system using machine learning techniques. In Advances in Communication, Devices and Networking (pp. 885-894). Springer, Singapore.
4. Elhag, S., Fernández, A., Bawakid, A., Alshomrani, S., & Herrera, F. (2015). On the combination of genetic fuzzy systems and pairwise learning for improving detection rates on intrusion detection systems. Expert Systems with Applications, 42(1), 193-202.
5. Guo, Y., Zhang, H., Zhang, L., Fang, L., & Li, F. (2019). A game theoretic approach to cooperative intrusion detection. Journal of computational science, 30, 118-126.
6. Kim, G., Lee, S., & Kim, S. (2014). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. Expert Systems with Applications, 41(4), 1690-1700.
7. Kumar, U., & Gohil, B. N. (2015). A survey on intrusion detection systems for cloud computing environment. International Journal of Computer Applications, 109(1).
8. Meena, G., & Choudhary, R. R. (2017, July). A review paper on IDS classification using KDD 99 and NSL KDD dataset in WEKA. In 2017 International Conference on Computer, Communications and Electronics (Comptelix) (pp. 553-558). IEEE.
9. Mohammadi, S., Mirvaziri, H., Ghazizadeh-Ahsaee, M., & Karimipour, H. (2019). Cyber intrusion detection by combined feature selection algorithm. Journal of information security and applications, 44, 80-88.
10. Nandeshwar, K. K., & Bijwe, K. B. (2015). Distributed Intrusion Detection System Using Mobile Agent Technology. International Journal of Engineering Research and General Science, 3(2).
11. Nskh, P., Varma, M. N., & Naik, R. R. (2016, May). Principle component analysis based intrusion detection system using support vector machine. In 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT) (pp. 1344-1350). IEEE.
12. Ranjan, R., & Sahoo, G. (2014). A new clustering approach for anomaly intrusion detection. arXiv preprint arXiv:1404.2772.
13. Salo, F., Nassif, A. B., & Essex, A. (2019). Dimensionality reduction with IG-PCA and ensemble classifier for network intrusion detection. Computer Networks, 148, 164-175.
14. Yahalom, R., Steren, A., Nameri, Y., Roytman, M., Porgador, A., & Elovici, Y. (2019). Improving the effectiveness of intrusion detection systems for hierarchical data. Knowledge-Based Systems, 168, 59-69.